DOCUMENT RESUME

ED 081 016 CS 200 678

AUTHOR O'Donnell, Roy C.
TITLE Some Syntactic Characteristics of Spoken and Written Discourse. Studies in Language Education, Report No. 4.
INSTITUTION Georgia Univ., Athens. Dept. of Language Education.
PUB DATE Sep 73
NOTE 18p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Composition (Literary); *Discourse Analysis; Graduate Students; Journalism; *Oral English; Professional Recognition; *Structural Analysis; Syntax

ABSTRACT

This study investigated the written and oral discourse of a subject with an established reputation as an author, editor, lecturer, and television host, to determine whether his written discourse was more complex syntactically than his spoken discourse. The oral language sample was taken from a taped television program in which the speaker responded to questions asked by three guest journalists; the written sample was taken from four newspaper columns written by the speaker. The samples were then segmented into minimal terminal syntactic units and the words in each sample were counted, as were the clauses, nonfinite verbals, passive constructions, auxiliaries, and attributive adjectives. Findings revealed (1) the median length of syntactic units was significantly greater in the written language samples; (2) the proportion of syntactic units containing one or more dependent clauses was significantly greater in written samples; (3) oral samples did not show as high a frequency of gerunds, participles, attributive adjectives, passive constructions and modal and perfective auxiliaries; and (4) noun clauses, infinitives, and progressive auxiliaries were observed more often in spoken discourse. (HOD)

FILMED FROM BEST AVAILABLE CO

ED 081016

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRO
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRE
SENT OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

SOME SYNTACTIC CHARACTERISTICS OF SPOKEN AND WRITTEN DISCOURSE

by

Roy C. O'Donnell

Studies in Language Education, Report No. 4
Department of Language Education, The University of Georgia
Athens, Georgia

September, 1973

25 800 678

TABLE OF CONTENTS

	page
Table of Contents	i
Tables	ii
Background and Statement of Problem	1
Related Studies	1
Rationale for Further Investigation	4
Statement of Problem	5
Description of Samples and Procedures	6
The Language Samples	6
Analytic Procedures	6
Findings and Conclusions	7
Findings	7
Conclusions	13
References	15

TABLES

page

Table I --Totals of Selected Variables Observed in 100 Syntactic Units of Spoken Discourse and in 100 Units of Written Discourse 7

Table II --Measures of Central Tendency and Variability of Words per Syntactic Unit in 100 Units of Spoken and 100 Units of Written Discourse 8

Table III--Comparison of Number of "Short" Syntactic Units with Number of "Long" Units in 100 Units of Spoken and 100 Units of Written Discourse. 9

Table IV --Comparison of Single-Clause and Multi-Clause Syntactic Units in 100 Units of Spoken and 100 Units of Written Discourse 10

Table V --Number of Clauses per Syntactic Unit in 100 Units of Spoken and 100 Units of Written Discourse 10

Table VI --Dependent Clauses in 100 Syntactic Units of Spoken Discourse and 100 Units of Written Discourse 11

Table VII--Non-Finite Verbals, Passives, Auxiliaries, Attributive Adjectives, and Mean Clause Length in 100 Syntactic Units of Spoken and 100 Units of Written Discourse 12

Teachers who tell their students that writing is simply "talk written down" emphasize the similarity of spoken and written discourse while ignoring the differences. No doubt the scarcity of useful information about the differences between these two modes of language makes it easier to ignore the differences than to make students purposefully aware of their existence. But while it is certainly true that speech and writing share a common grammatical system, it appears to be just as true that skilled writing exhibits distinctive characteristics which students should be able to recognize and use effectively. To gather more information about the syntactic characteristics that distinguish written from spoken discourse is the purpose of this study.

Background and Statement of Problem

Related Studies

More than fifty years ago, Woolbert (1922) called attention to the similarities and differences between spoken and written language. He said: "Speaking and writing are alike--and different. Just how like and how different has never been adequately stated." Less than a decade ago, Gleason (1965) made the following comment concerning the differences; "...we know that written and spoken structures differ to some extent, and that there are differences in the underlying grammatical structure, but we have no accurate idea of their magnitude." Gibson, et. al. (1966)

commented as follows concerning the state of knowledge about differences between speaking and writing:

> Numerous scholars have initiated empirical studies to describe and explain the character of the supposed variations in written and spoken messages. Their research has focused upon such matters as the types of words used, sentence length, grammatical characteristics, overall comprehensibility, and modes of expression. However, the findings in the empirical research relative to similarities and differences in written and spoken messages have been contradictory and unclear at best, resulting in no empirically demonstrated description of the similarities and differences between written and spoken messages.

In an attempt to remedy the deficiency in knowledge, Gibson and his colleagues compared samples of speech and writing of forty-five beginning speech students at two universities. They computed Flesch Reading Ease scores (including average sentence length and average number of syllables per 100 words), Flesch Human Interest scores, and Type-Token Ratios (TTR). They found fewer words per sentence, fewer syllables per word, and less diversity of vocabulary in spoken than in written discourse; interest scores were higher in speech than in writing.

These findings are not essentially different from those of an earlier study conducted in Holland by Drieman (1962). He found that written language as compared to spoken has shorter texts (greater conciseness), longer words, fewer words of one syllable, more attributive adjectives, and a more varied vocabulary.

Two studies reported by Devito (1966, 1967) searched for more precise differences between oral and written discourse in vocabulary and in level of abstraction. He found in language samples from ten university faculty members that oral discourse differs from written in the following respects: more self-reference words, more pseudo-quantifying words (<u>much</u>, <u>many</u>, etc.), more allness terms (<u>all</u>, <u>none</u>, etc.), more qualification

terms (if, but, etc.), and more terms indicative of consciousness of projection (seems, appears, etc.). Applying the Gillie Level of Abstraction formula, Devito found that oral language is less abstract than written; it has more finite verbs and fewer nouns of abstraction.

Harrell (1957) made grammatical analyses and comparisons of written and oral narratives of 320 children from nine to fifteen years of age. The children were shown a short movie and asked to write the story of the movie; later they were shown another movie and asked to tell the story privately to the investigator. He found that the children used more subordinate clauses in writing than in speaking, and that this difference increased with age. In greater detail, Harrell found that the children used more adverb and adjective clauses in writing than in speaking, but that they used more noun clauses and a greater percentage of all types of adverbial clauses, except clauses of time and cause, in their oral stories.

Another study comparing grammatical characteristics of speech and writing was made by Blankenship (1962). Utilizing a modification of Fries' grammatical system, she compared four speeches delivered on a university campus with published articles written by the speakers. Blankenship found little variation in "sentence" length, some variation in sentence patterns, and some differences in use of word classes in various positions. She also found more use of transitive verbs in oral language and more use of passive constructions in written language. Recognizing the lack of findings of clear-cut grammatical distinctions in her language samples, Blankenship said: "This study indicates that syntactical structure is determined by an individual's style rather than by read/heard purpose."

Rationale for Further Investigation

In retrospect, it seems likely that limitations in the analytic system used by Blankenship could account for the lack of clear differences between speech and writing in her study. Thus, it seems unwarranted to rule out the possibility of characteristic syntactic differences between oral and written discourse, but it seems highly desirable to look more closely at individual style as a factor in determining the differences. There is clearly a need for further comparison of syntax in oral and written discourse, utilizing a more efficient system of analysis and focusing on the language used by an individual.

The approach to syntactic analysis employed by O'Donnell, Griffin, and Norris (1967) in a study of children's language development appears to be potentially useful for further comparison of spoken and written language structures. In their study of children's oral and written language in the middle grades they used techniques of analysis derived in part from Chomsky (1957, 1965) and a basic syntactic unit identified by Hunt (1965). This minimal terminable syntactic unit (T-unit) contains one independent clause and the dependent clauses (if any) related to it. The length of this basic syntactic unit reflects the number of generalized (sentence-combining) transformations applied to its underlying structure. In terms of traditional grammar, the number of words in the unit reflects, among other things, the number of dependent clauses, non-finite verbals, and modifiers of nouns.

O'Donnell and his colleagues found that the mean length of syntactic units was significantly greater in oral than in written language samples of third graders, but not for children in grades five and seven. In these grades the mean number of sentence-combining transformations per syntactic

unit was significantly greater in writing than in speech. More specifically, constructions containing modifiers of nouns, coordinate predicates, and adverbial clauses appeared more frequently in written than in oral discourse of children in grades five and seven. These phenomena were interpreted as evidence of greater syntactic complexity in written discourse than in oral.

Statement of Problem

The present study is designed to investigate the question of whether written discourse of an individual is more complex syntactically than his spoken discourse. Since it can be demonstrated that length of minimal terminable syntactic unit generally reflects the degree of syntactic complexity, the major hypothesis tested is that written discourse has a greater proportion of "long" syntactic units than does spoken discourse("long" units are defined as those greater in length than the median length of spoken and written units combined). The Median Test is used to examine this hypothesis. Since length of syntactic units is partially accounted for by the presence of dependent clauses, a related hypothesis tested in this study is that written discourse has a greater proportion of syntactic units containing dependent clauses than does spoken. This secondary hypothesis is tested by means of Chi Square calculation. Other selected syntactic variables are examined but not tested for statistical significance.

Description of Samples and Procedures

The Language Samples

The samples of spoken and written discourse analyzed in this study were produced by an adult male, who is a university graduate and who has an established reputation as an author, editor, lecturer, and television host. The sample of oral language is taken from a television program taped in January 1973, in which the speaker responded to questions on a variety of topics asked by three guest journalists. The written sample was taken from four newspaper columns written by the speaker and published shortly after the telecast. These samples are similar in that they represent an individual's public expression of ideas and opinions on a variety of topics of public interest and concern. Another important aspect of their similarity is that they were produced by a mature person who has demonstrated adeptness in both spoken and written discourse.

Analytic Procedures

Copies of spoken and written discourse were segmented into minimal terminable syntactic units (T-units). One hundred units of speech, excluding responses of less than 100 words to assure samples of continuous discourse; and 100 units of writing, excluding material quoted from other writers, constitutedthe language samples analyzed. The total number of words in each sample was counted, as was the number of clauses, non-finite verbals, passive constructions, auxiliaries, and attributive adjectives. The resulting data were processed by appropriate statistical computations and analyses to see if the spoken and written language samples differed in respect to certain syntactic characteristics. Findings are reported in Tables I–VII and interpreted in the following paragraphs.

Findings and Conclusions

Findings

Data presented in Table I indicate that there are more words per 100 syntactic units of written as compared to oral discourse and more dependent clauses. On a percentage basis the spoken units have only 72% as many words and 84% as many dependent clauses as the written units have.

TABLE I

Totals of Selected Variables Observed in 100 Syntactic Units of Spoken Discourse and in 100 Units of Written Discourse

	SPOKEN	WRITTEN
Total Words	1792	2497
Total Clauses	215	237
Total Dependent Clauses	115	137

Table II presents data showing that syntactic units in writing exhibit greater average length and greater variability in length than units in speech. Although a statistical test would show that the mean difference of about seven words in length of syntactic units in speech and writing is significant, a more appropriate test for comparison of number of words per unit is the Median Test. The Median Test is actually a Chi Square calculation using the combined median (18.5 words in these language samples) as a dividing point to form frequency categories.

TABLE II

Measures of Central Tendency and Variability of Words per Syntactic Unit in 100 Units of Spoken and 100 Units of Written Discourse

	SPOKEN	WRITTEN
Median	13	22
Range	54 (3-56)	64 (3-66)
Mean	17.92	24.97
Standard Deviation	12.60	14.95

The number of syntactic units shorter than the combined median ("short" units) and the number longer than the combined median ("long" units) are shown in Table III. Since each sample consists of exactly 100 units, the number of units in each instance can be read as a percentage. It is interesting to note that the proportions of "long" and "short" syntactic units are exactly reversed in these samples of oral and written discourse. The Chi Square value of 11.52 is significant below the .01 level, indicating that the median length of written syntactic units is greater than that of spoken units. In other words, the proportion of "short" syntactic units is significantly greater in speech than in writing and the proportion of "long" syntactic units is greater in writing than in speech. Therefore, the major hypothesis of this study is sustained.

TABLE III

Comparison of Number of "Short" Syntactic Units with Number of "Long" Units in 100 Units of Spoken and 100 Units of Written Discourse

	SPOKEN	WRITTEN
Number of short units (18 words or less)	62	38
Number of long units (19 words or more)	38	62

$\chi^2=11.52$, $p<.01$

Since the minimal terminable syntactic unit used in this study consists of one independent clause and any dependent clauses accompanying it, and since the presence of dependent clauses contributes to syntactic complexity, the comparison of speech and writing with respect to the number of single and multi-clause units shown in Table IV is revealing. Forty-four per cent of the spoken units consist of one independent clause with no dependent clauses, whereas only 32% of the written units are thus constituted. Conversely 68% of the written units and 56% of the spoken units have one or more dependent clauses accompanying the independent clause. The Chi Square test (one-tailed) shows that the proportion of single clause units is significantly greater ($p < .05$) in spoken than in written discourse. Thus, the secondary hypothesis of this study is also sustained. Stated another way, there are more syntactic units containing one or more dependent clauses in writing than in speech. Hence the written units tend to be more complex in this aspect of syntactic structure.

TABLE IV

Comparison of Single-Clause and Multi-Clause Syntactic Units in 100 Units of Spoken and 100 Units of Written Discourse

	SPOKEN	WRITTEN
Syntactic Units with No Dependent Clauses	44	32
Syntactic Units with One or More Dependent Clauses	56	68

$\chi^2 = 3.05$, $p < .05$ on one-tailed test

A more detailed break-down of the number of clauses composing the syntactic units is given in Table V.

TABLE V

Number of Clauses per Syntactic Unit in 100 Units of Spoken and 100 Units of Written Discourse

	SPOKEN	WRITTEN
Syntactic units consisting of:		
1 clause	44	32
2 clauses	25	36
3 clauses	15	12
4 clauses	10	10
5 clauses	4	4
6 clauses	1	5
7 clauses	0	1
8 clauses	1	0
Mean Number of clauses per unit	2.15	2.37

TABLE VI

Dependent Clauses in 100 Syntactic Units of Spoken Discourse and 100 Units of Written Discourse

	SPOKEN	WRITTEN
Nominal Clauses	57	53
Adjectival Clauses	30	43
Adverbial Clauses	22	33
Interjected Clauses	6	8
Total Dependent Clauses	115	137

In Table VI, the data show that the nominal clause is the most frequently occurring type of dependent clause and that it occurs more often in oral than in written discourse. All other kinds of dependent clauses, however, are less frequent in speech than in writing. Spoken units have only 70% as many adjectival, 67% as many adverbial, and 75% as many interjected clauses.

Because clause length is accounted for by a variety of syntactic characteristics, it seemed worthwhile to compare the two samples in respect to frequency of occurrence of passives, infinitives, gerunds, participles, auxiliaries, and attributive adjectives.

Table VII shows that the most frequently occurring non-finite verbal is the infinitive and that it occurs more often in spoken than in written discourse. Since gerunds occur three times as often and participles appear twice as often in written as in oral, however, the total number of non-finite verbals is greater in writing than in speech. Passives are not used often in either sample, but they are slightly more than twice as frequent in

writing. The total number of auxiliaries is greater in writing than in speech, with progressive auxiliaries (forms of be with verbs) occurring slightly more often in spoken discourse. The number of perfective auxiliaries (forms of have with verbs) in oral is only 63% as great as in written discourse, and the number of modal auxiliaries only 68%. Spoken units have 73% as many attributive adjectives as written. As a result of these and other syntactic phenomena, the mean number of words per clause in speech is only 79% as great as that in writing.

TABLE VII

Non-Finite Verbals, Passives, Auxiliaries, Attributive Adjectives, and Mean Clause Length in 100 Syntactic Units of Spoken and 100 Units of Written Discourse

	SPOKEN	WRITTEN
Total Non-Finite Verbals	80	97
Infinitives	61	52
Gerunds	7	21
Participles	12	24
Total Passive Constructions	14	29
Total Auxiliaries	40	53
Modals	17	25
Perfectives	12	19
Progressives	11	9
Total Attributive Adjectives	74	101
Mean Number of Words per Clause	8.33	10.54

In summary, the findings of this study of syntactic characteristics of spoken and written discourse are stated as follows:

1) The median length of syntactic units is significantly greater in the written language samples than in the spoken samples. In other words, the proportion of "long" units is greater in written than in spoken language samples. This greater average length of written syntactic units is accounted for in part by the phenomena observed in other findings of the study.
2) The proportion of syntactic units containing one or more dependent clauses is significantly greater in the samples of writing than in the samples of speech. This phenomenon is reflected in the greater average number of dependent clauses in the written units.
3) Gerunds, participles, attributive adjectives, passive constructions, and modal and perfective auxiliaries occur more often in written samples than in oral samples. These phenomena partially account for the greater average length of written clauses.
4) Only noun clauses, infinitives, and progressive auxiliaries were observed more often in spoken than in written discourse.

Conclusions

Since these samples were taken from one individual's oral and written language, no assured inferences can be made about the syntactic differences between speech and writing in general. In view of the limited size of the samples, it would also be unwise to make any sweeping generalizations about the spoken and written syntax of an individual.

However, the findings summarized above support the conclusion that in the samples of spoken and written discourse analyzed in this study, writing has a greater degree of syntactic complexity or syntactic density than speech, a condition associated with compactness in verbal organization and heavy information load. In this, they complement some of the less definitely supported findings of previous studies of the syntactic characteristics of speech and writing in general. Thus, these findings make it possible for teachers to offer, with a somewhat greater degree of confidence, information to students concerning the characteristics that distinguish skilled writing from spoken discourse.

Also, the clearly delineated differences in these language samples with respect to greater length of syntactic units, greater number of dependent clauses, greater number of passive constructions, greater number of attributive adjectives, greater number of non-finite verbals, and greater number of auxiliaries in written discourse is impressive and suggests the potential value of the system of analysis used in this study for other studies of syntactic differences between spoken and written discourse; for it can still be said that "we have no accurate idea of their magnitude."

REFERENCES

Blankenship, Jane. 1962. "A Linguistic Analysis of Oral and Written Style," QUARTERLY JOURNAL OF SPEECH, 48, 419-422.

Chomsky, Noam A. 1957. SYNTACTIC STRUCTURES. 's Gravenhage: Mouton and Co.

Chomsky, Noam A. 1965. ASPECTS OF THE THEORY OF SYNTAX. Cambridge, Mass.: MIT Press.

Devito, J. A. 1966. Psychogrammatical Factors in Oral and Written Discourse by Skilled Communicators," SPEECH MONOGRAPHS, 33, 73-76.

Devito, J. A. 1967, "Levels of Abstraction in Spoken and Written Language," JOURNAL OF COMMUNICATION, 17, 354-361.

Drieman, G. H. J. 1962. "Differences Between Written and Spoken Languages: an Exploratory Study," ACTA PSYCHOLOGICA, 20, 36-57 and 78-100.

Gibson, J. W., C. R. Gruner, R. J. Kibler, and F. Kelly. 1966. "A Quantitative Examination of Differences and Similarities in Written and Spoken Messages." SPEECH MONOGRAPHS, 33, 444-451.

Gleason, H. A., Jr., 1965. LINGUISTICS AND ENGLISH GRAMMAR. New York: Holt, Rinehart and Winston.

Harrell, Lester E., Jr., 1957. "A Comparison of Oral and Written Language in School-Age Children," MONOGRAPHS OF THE SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT, 22:3.

Hunt, Kellogg W. 1965. GRAMMATICAL STRUCTURES WRITTEN AT THREE GRADE LEVELS. NCTE Research Report No. 3. Champaign, Ill.: National Council of Teachers of English.

O'Donnell, Roy C., William J. Griffin, and Raymond C. Norris. 1967. "A Transformational Analysis of Oral and Written Grammatical Structures in the Language of Children in Grades Three, Five, and Seven," THE JOURNAL OF EDUCATIONAL RESEARCH, 61,36-39.

Woolbert, C. H. 1922. "Speaking and Writing--A Study of Differences," QUARTERLY JOURNAL OF SPEECH EDUCATION, 8, 271-275.